Monotone Transformationen

Markus J. Stroppel

1. November 2019

Gegeben sind zwei (Eingangs-)Werte $x_{\min}$ und $x_{\max}$, dazu zwei (Ausgangs-)Werte $y_{\min}$ und $y_{\max}$. Wir suchen eine möglichst einfache Transformation T, die $x_{\min}$ auf $y_{\min}$ und $x_{\max}$ auf $y_{\max}$ abbildet (und alle Werte dazwischen „passend“: auf jeden Fall so, dass aus $x_1 < x_2$ auch $T(x_1) < T(x_2)$ folgt).

1 Affine Transformation

Die einfachste Möglichkeit ist die *affine Transformation*: Man sucht sich eine Steigung a und eine Verschiebung b so, dass

$$\begin{aligned} a \cdot x_{\min} + b &= y_{\min} \\ a \cdot x_{\max} + b &= y_{\max} \,. \end{aligned}$$

Dann setzt man $T(x) := a \cdot x + b$.

Um a und b zu bestimmen, ziehen wir die erste Gleichung von der zweiten ab und erhalten

$$a \cdot (x_{\max} - x_{\min}) = y_{\max} - y_{\min} \,.$$

Daraus kann man den Faktor a ausrechnen als

$$a = \frac{y_{\max} - y_{\min}}{x_{\max} - x_{\min}}$$

und dann aus der ersten Gleichung die Verschiebung

$$b = y_{\min} - a \cdot x_{\min} = y_{\min} - \frac{y_{\max} - y_{\min}}{x_{\max} - x_{\min}} \cdot x_{\min} = \frac{y_{\min} \cdot x_{\max} - y_{\max} \cdot x_{\min}}{x_{\max} - x_{\min}} \,.$$

Für die konkreten Eingangs- und Ausgangswerte (siehe Abb. 1)

$$x_{\min} = 0.1, \quad x_{\max} = 200\,000, \quad y_{\min} = 1.0, \quad y_{\max} = 20$$

ergibt sich $a = \frac{20-1}{200\,000-0.1} \approx \frac{19}{200\,000} = 0,000095 \approx \frac{1}{10\,000}$ und dann $b = \frac{1 \cdot 200\,000 - 20 \cdot 0.1}{200\,000-0.1} \approx 1$; in der Tat:

$$\begin{aligned} \tfrac{19}{200\,000} \cdot 0.1 \quad &+ 1 \approx 1 \,, \\ \tfrac{19}{200\,000} \cdot 200\,000 &+ 1 = 20 \,. \end{aligned}$$

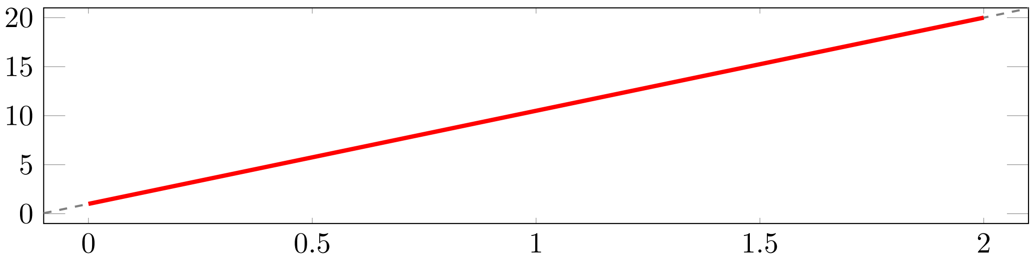


Abbildung 1: Lineare Transformation: interpoliert durch eine Gerade

2 Transformation durch interpolierende Parabeln

Manchmal will man die transformierte Größenskala an verschiedenen Stellen unterschiedlich stark spreizen. Der rechnerisch einfachste Weg ist über eine interpolierende Parabel (an Stelle der interpolierenden Geraden bei der affinen Transformation).

Parabeln erhält man als Graphen von Funktionen p mit Funktionsterm $p(x) = a \cdot x^2 + b \cdot x + c$. Wenn man (wie im vorliegenden Anwendungsfall) kontrollieren will, dass die transformierende Funktion im Intervall von x_{min} bis x_{max} durchgehend ansteigt, sollte man den Scheitel der Parabel kennen. Deswegen benutzt man besser die Darstellung

$$p(x) = a \cdot (x - s)^2 + d\,,$$

wobei sich s und d aus a, b und c ergeben durch den Vergleich $a \cdot x^2 + b \cdot x + c = p(x) = a \cdot (x - s)^2 + d = a \cdot x^2 - 2 \cdot a \cdot s \cdot x + a \cdot s^2 + d$ als $s = \frac{b}{2 \cdot a}$ und $d = c - a \cdot s^2$.

Damit die Funktion p im fraglichen Bereich ansteigt, muss entweder $s \leq x_{\text{min}}$ und $a > 0$ oder aber $x_{\text{max}} \leq s$ und $a < 0$ sein; im ersten Fall ist die Parabel nach oben, im zweiten nach unten geöffnet.

Wir wollen noch $p(x_{\text{min}}) = y_{\text{min}}$ und $p(x_{\text{max}}) = y_{\text{max}}$ sichern. Das geschieht durch Lösen des Gleichungssystems

$$\begin{aligned} a \cdot (x_{\text{min}} - s)^2 + d &= y_{\text{min}} \\ a \cdot (x_{\text{max}} - s)^2 + d &= y_{\text{max}}\,. \end{aligned}$$

Subtraktion der Gleichungen ergibt jetzt

$$a \cdot \left((x_{\text{max}} - s)^2 - (x_{\text{min}} - s)^2\right) = y_{\text{max}} - y_{\text{min}}\,,$$

also

$$a = \frac{y_{\text{max}} - y_{\text{min}}}{(x_{\text{max}} - s)^2 - (x_{\text{min}} - s)^2}$$

und damit dann

$$\begin{aligned} d = y_{\text{min}} - a \cdot (x_{\text{min}} - s)^2 &= y_{\text{min}} - \frac{(y_{\text{max}} - y_{\text{min}})((x_{\text{min}} - s)^2)}{(x_{\text{max}} - s)^2 - (x_{\text{min}} - s)^2} \\ &= \frac{y_{\text{min}}(x_{\text{max}} - s)^2 - y_{\text{max}}(x_{\text{min}} - s)^2}{(x_{\text{max}} - s)^2 - (x_{\text{min}} - s)^2}\,. \end{aligned}$$

Wir haben jetzt den Parameter s noch frei wählbar (unter Berücksichtigung der Bedingung $s \le x_{\text{min}}$ bzw. $x_{\text{max}} \le s$. Damit ergeben sich als mögliche Transformationen die Funktionen aus der durch $s \in (-\infty, x_{\text{min}}] \cup [x_{\text{max}}, +\infty)$ parametrisierten Schar

$$\begin{aligned} T_s(x) &:= \frac{(y_{\text{max}} - y_{\text{min}})(x-s)^2 + y_{\text{min}}(x_{\text{max}} - s)^2 - y_{\text{max}}(x_{\text{min}} - s)^2}{(x_{\text{max}} - s)^2 - (x_{\text{min}} - s)^2} \\ &= \frac{(y_{\text{max}} - y_{\text{min}})(x^2 - 2\,s\,x) + y_{\text{min}}(x_{\text{max}} - 2\,s\,x_{\text{max}}) - y_{\text{max}}(x_{\text{min}} - 2\,s\,x_{\text{min}})}{x_{\text{max}}^2 - x_{\text{min}}^2 - 2\,s\,(x_{\text{max}} - x_{\text{min}})}\,. \end{aligned}$$

Durch geeignete Wahl des Scharparameters s (siehe Abb. 2, Abb. 3) kann man erreichen, dass Werte knapp unter dem größten sich deutlicher vom größten unterscheiden ($s \le x_{\text{min}}$) oder analog beim kleinsten ($x_{\text{max}} < s$).

Konkrete Beispiele werden gleich berechnet und skizziert (siehe Abb. 2, Abb. 3) für die folgenden Eingangs- und Ausgangswerte:

$$\begin{aligned} x_{\text{min}} &= 1 \qquad & x_{\text{max}} &= 4 \\ y_{\text{min}} &= 3 \qquad & y_{\text{max}} &= 5 \end{aligned}$$

Es ergibt sich (noch abhängig von s):

$$\begin{aligned} a &= \frac{5-3}{(4-s)^2 - (1-s)^2} \\ d &= \frac{3(4-s)^2 - 5(1-s)^2}{(4-s)^2 - (1-s)^2}\,. \end{aligned}$$

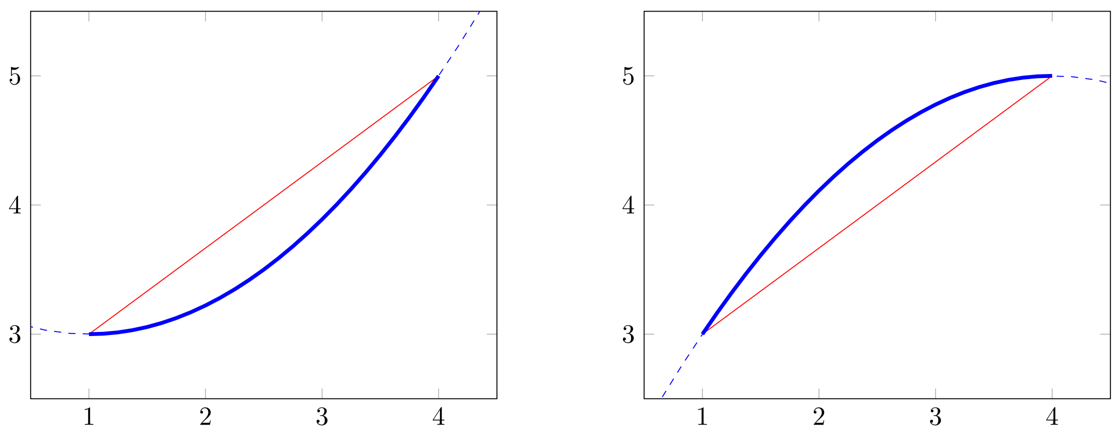
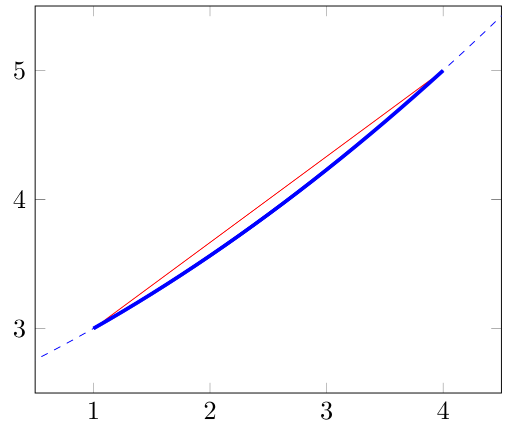


Abbildung 2: interpoliert durch eine Parabel: Scheitel $s = x_{\text{min}} = 1$ (links) bzw. $s = x_{\text{max}} = 4$ (rechts); zum Vergleich die lineare Interpolation (rot)

Speziell für $s = 1$:

$$\begin{aligned} a &= \frac{5-3}{(4-1)^2 - (1-1)^2} = \frac{2}{9} \\ d &= \frac{3(4-1)^2 - 5(1-1)^2}{(4-1)^2 - (1-1)^2} = 3\,. \end{aligned}$$



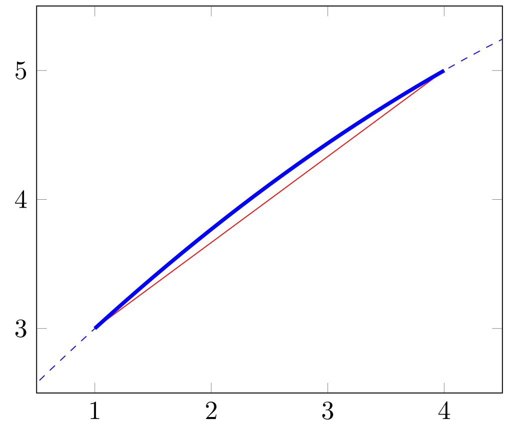


Abbildung 3: interpoliert durch eine Parabel: Scheitel $s < x_{\text{min}} = 1$ (links) bzw. $s > x_{\text{max}} = 4$ (rechts); zum Vergleich die lineare Interpolation (rot)

Je weiter der Scheitel s unterhalb von x_{min} oder oberhalb von x_{max} ist, desto mehr nähert sich der Parabelbogen der Strecke an, die man als Schaubild der linearen Interpolation bekommt (siehe Abb. 2 im Vergleich zu Abb. 3).

Das sieht man so ein:

Wir erweitern den vorher erhaltenen Ausdruck für $T_s(x)$ mit $\frac{1}{s}$ und erhalten

$$\begin{aligned} T_s(x) &= \frac{\frac{1}{s}\left((y_{\text{max}} - y_{\text{min}})(x^2 - 2\,s\,x) + y_{\text{min}}(x_{\text{max}} - 2\,s\,x_{\text{max}}) - y_{\text{max}}(x_{\text{min}} - 2\,s\,x_{\text{min}})\right)}{\frac{1}{s}\left(x_{\text{max}}^2 - x_{\text{min}}^2 - 2\,s\,(x_{\text{max}} - x_{\text{min}})\right)} \\ &= \frac{(y_{\text{max}} - y_{\text{min}})(\frac{1}{s}x^2 - 2\,x) + y_{\text{min}}(\frac{1}{s}x_{\text{max}} - 2\,x_{\text{max}}) - y_{\text{max}}(\frac{1}{s}x_{\text{min}} - 2\,x_{\text{min}})}{\frac{1}{s}(x_{\text{max}}^2 - x_{\text{min}}^2) - 2\,(x_{\text{max}} - x_{\text{min}})} \end{aligned}$$

Wenn s gegen $+\infty$ oder gegen $-\infty$ strebt, strebt deswegen $T_s(x)$ gegen

$$\begin{aligned} T(x) &= \frac{(y_{\text{max}} - y_{\text{min}})(-2\,x) + y_{\text{min}}(-2\,x_{\text{max}}) - y_{\text{max}}(-2\,x_{\text{min}})}{-2\,(x_{\text{max}} - x_{\text{min}})} \\ &= \frac{y_{\text{max}} - y_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}x + \frac{y_{\text{min}}\,x_{\text{max}} - y_{\text{max}}\,x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}\,. \end{aligned}$$

Das ist die lineare Interpolation aus Abschnitt 1

3 Noch höherer Grad

Man kann auch mit Polynomen der Form $p(x) = a \cdot (x - s)^4 + d$ interpolieren. Das sind ziemlich spezielle Polynome vom Grad 4; die immer noch einen Scheitel haben (bei s). Wie im vorigen Abschnitt wählen wir entweder $s \leq x_{\text{min}}$ oder $s \geq x_{\text{max}}$, um zu erreichen, dass die interpolierende Funktion im relevanten Intervall ansteigt.

Die Formeln von vorhin übertragen sich; wir erhalten

$$a = \frac{y_{\text{max}} - y_{\text{min}}}{(x_{\text{max}} - s)^4 - (x_{\text{min}} - s)^4}$$

und damit dann

$$d = y_{\text{min}} - a \cdot (x_{\text{min}} - s)^4 = \frac{y_{\text{min}}(x_{\text{max}} - s)^4 - y_{\text{max}}(x_{\text{min}} - s)^4}{(x_{\text{max}} - s)^4 - (x_{\text{min}} - s)^4}\,.$$

Der Graph der interpolierenden Funktion wird jetzt am einen Ende viel steiler als die lineare Interpolation; am anderen Ende dagegen extrem flach (siehe Abb. 4). Durch Verschieben des Scheitels s unter x_{min} oder über x_{max} hinaus kann man diesen Effekt wieder abmildern (siehe Abb. 5).

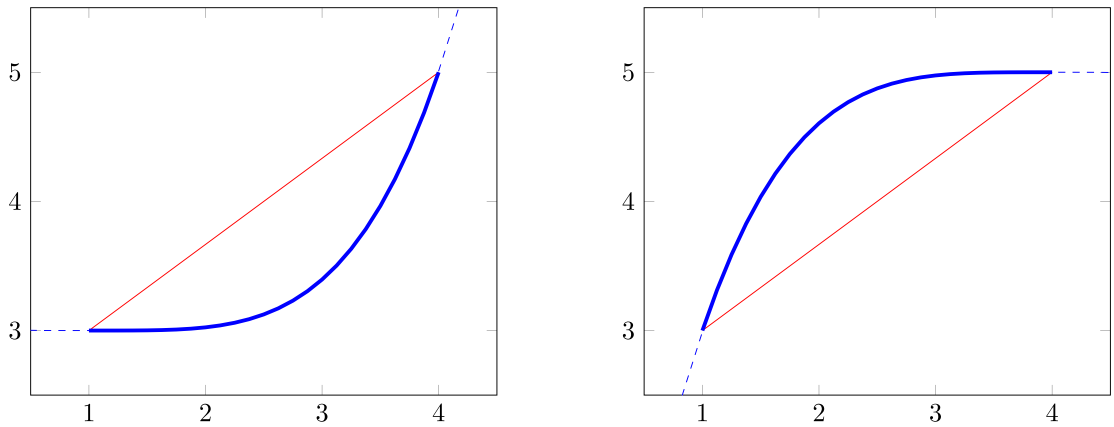


Abbildung 4: interpoliert durch ein Polynom der Form $a \cdot (x - s)^4 + d$: Scheitel $s = x_{\text{min}} = 1$ (links) bzw. $s = x_{\text{max}} = 4$ (rechts); zum Vergleich die lineare Interpolation (rot)

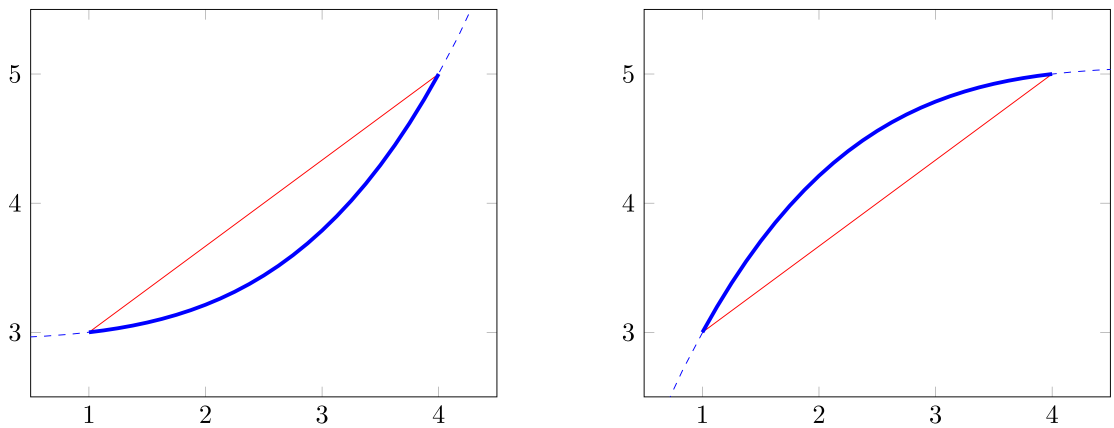


Abbildung 5: interpoliert durch ein Polynom der Form $a \cdot (x - s)^4 + d$: Scheitel $s = -1 < x_{\text{min}}$ (links) bzw. $s = 5 > x_{\text{max}}$ (rechts); zum Vergleich die lineare Interpolation (rot)

4 Spreizung an beiden Enden

Die Spreizung, die wir in Abschnitt 2 und stärker noch in Abschnitt 3 jeweils an einem (von uns wählbaren) Ende des zu transformierenden Intervalls erreicht haben, kann man durch Verwendung noch komplizierterer Funktionen auch an beiden Enden des Intervalls gleichzeitig erreichen.

Dazu verwendet man Funktionen, die an beiden Enden des Intervalls stärker steigen als die affine Transformation.

Wir beschränken uns hier auf zwei Beispiele:

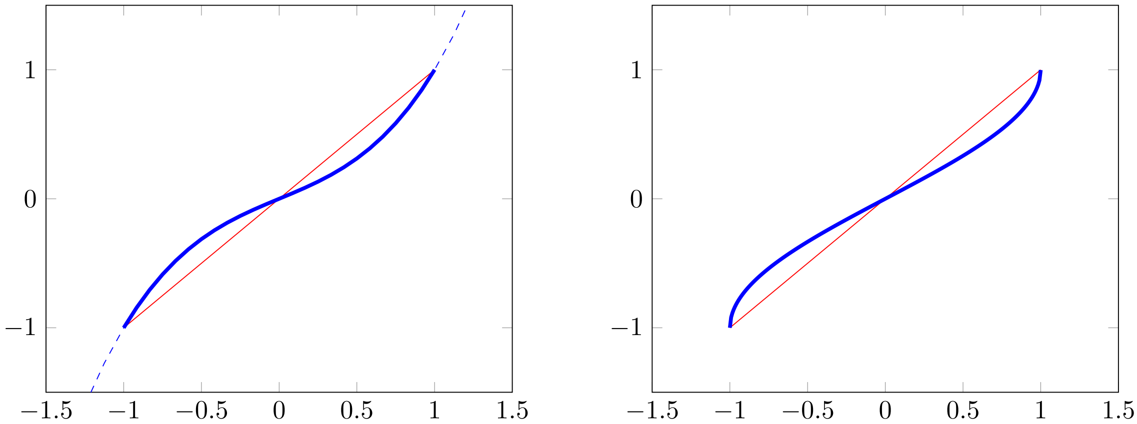


Abbildung 6: interpoliert durch ein Polynom vom Grad 3 (links) und durch die Umkehrfunktion $\sin^{-1}$ (rechts); zum Vergleich die lineare Interpolation (rot)

Die Umkehrfunktion der Sinus-Funktion wird an den Enden des Intervalls $[-1, 1]$ besonders steil (die Tangenten in den Endpunkten stehen vertikal).